Is this approach more representative of the engineering design of a network because it does not rely on census-mapping conventions? ¶ 47

Comment:

As we discussed earlier, the enhanced BCPM demonstrates that a network design can incorporate both census data and technological constraints. Furthermore, the Joint Sponsors submit that it is appropriate for a cost proxy model to capture both so that it reflects reality.

Could this proposal be incorporated into either BCPM or Hatfield and, if so, would any alterations in BCPM or Hatfield be necessary? ¶ 47

Comment:

This proposal has been incorporated into the enhanced BCPM.

c. **Line Count**

The model selected must assign lines to CBGs or wire centers. ¶ 48 The state members of the Joint Board voiced concern about the estimates of customer lines per wire center generated by each model. They assert that errors in these estimates might be traced to assignment of CBGs to incorrect wire centers. The Majority State Members Report calls for a requirement that models should match within 10% actual wire center line counts. ¶ 49

BCPM uses 1995 Census estimates of the number of households in each CBG. BCPM estimates the total number of residential lines for each CBG by allocating actual residential access lines in a state based on the number of households in a CBG. BCPM estimates the number of business lines by allocating actual business access lines in a state to each CBG based on the number of employees in the CBG per Dunn & Bradstreet data. Once lines have been allocated to the CBGs, BCPM assigns CBGs to wire centers by assigning the CBG to the wire center closest to the centroid of that CBG. ¶ 50

Starting from a 1995 Census household estimate, Hatfield estimates the residential line counts for each CBG. It removes households without telephones (according to 1990 Census information) and adds second lines for some households using an estimated relationship between second lines and CBG data about the income and age of consumers. Hatfield assigns business lines to CBGs on the basis of the number of employees within a CBG, as BCPM does, but also considers the relative intensity of telephone demand across different industries. The detailed analysis that underlies these assignments was not filed with the Commission. The sum of all residential and business lines assigned to CBGs matches state totals for residential and business lines. Each CBG is assigned to the ILEC wire center that serves more *customers in that CBG than any other. Hatfield attempts to include special access lines, but BCPM does not. ¶ 51*

Both models use a closure factor: a ratio of line counts, as provided by NECA and ARMIS databases, compared to the models' estimates, to adjust the estimates to reflect the actual ILEC line counts. Neither model discloses the closing factors for all lines that are used in their line count calculations. ¶ 52

The Commission concludes that neither Hatfield nor BCPM algorithms accurately predict line count. The FCC seeks comment on what changes can be made to those algorithms to improve their accuracy. ¶ 53

Comment:

The enhanced BCPM methodology will incorporate improved granularity in locating customers and in mapping wire center boundaries which the Joint Sponsors believe will greatly improve the accuracy in assigning lines to wire centers.

The Commission says that the models' algorithms should produce estimates that are accurate enough to avoid the need for a large closing factor to force the line-count estimate to match the wire center line count. The Commission tentatively concludes that the sizes and uses of models' closing factors should be evident to the user so that they may be evaluated. The Commission seeks comment on whether the model should adopt a maximum closing factor of 10%, as suggested by the state members of the Joint Board. ¶ 53

Comment:

Where actual and predicted line counts differ significantly, model sponsors should be required to document and justify why such discrepancy would occur. It is possible that discrepancies may be caused by a mismatch in the date of the census data and the date of the actual line count. Also, the method by which the Bureau of the Census updated 1990 census counts to 1995 projections could potentially pose a problem. Another factor which could contribute to discrepancy would be in the estimation of business lines, or the projection of additional residential lines, particularly with the large recent growth in second lines for Internet access.

Perhaps a more meaningful statistic would be to take the offices which would qualify for high-cost support and compare actual vs. predicted lines for these offices. Any carrier claiming support from the high-cost fund should be required by the funds Administrator to certify the number of customers within each designated funding area which it actually serves. Matching the number of customers currently served in each wire center, while an important indicator of model accuracy, should not be the sole means of evaluating the accuracy of a model in estimating forward-looking costs.

The Joint Sponsors have made significant improvements in the enhanced BCPM customer location algorithms which we believe will greatly improve the correlation of estimated vs. predicted line counts. Both model sponsors should be required to document the level of closure factor employed in their models.

The Commission seeks comment on whether other data sources could be used to enhance the models' algorithms or be used to create an alternative method for determining line counts. ¶ 53

Comment:

The Joint Sponsors believe that the additional data sources used in constructing the enhanced BCPM customer location algorithms will significantly improve the accuracy of the proxy models.

The Commission seeks comment on whether it should assign business lines to geographic units by using commercially produced maps that give the coordinates of all businesses located in the U.S. along with their employment by standard SIC code. ¶ 53

Comment:

The Joint Sponsors believe this type of data can significantly improve the proxy models. The Joint Sponsors plan to use PNR data that uses the SIC in algorithms to develop business line counts as well as business location data.

The Commission seeks comment on whether such a method should use some multiple of the employment data to estimate the number of business lines in each grid block. ¶ 53

Comment:

See above.

In the alternative, the Commission seeks comment on whether there are any databases that use zip code information or geo-coding information that could be used to improve the line-count estimation process. ¶ 53

Comment:

If such a data source were publicly available, the Joint Sponsors believe it could significantly improve the proxy models. At this time, however, the Joint Sponsors are not aware of such a data source. Whether or not this data exists, however, the Joint Sponsors submit that it is not appropriate to engineer on a customer by customer basis. Rather, a model should build to a collection of customers. Therefore, whatever is the source of customer locations (e.g., Census Block, Geocoded customer points), some method must be employed to group these customers into typical engineering areas. The Joint Sponsors believe the new BCPM Grid concept is the ideal tool in this regard.

IV. CONCLUSION

By enhancing the customer location platform design in the manner described above, the Joint Sponsors believe that BCPM now provides superior outputs on which to base decisions regarding the high-cost support fund.

Respectfully submitted,

SPRINT LOCAL TELEPHONE
COMPANIES

By: Jay C. Keithley

Jay C. Keithley
Sandra K. Williams
1850 M Street
Suite 1110
Washington, D.C. 20036
(202) 828-7453

Attorneys for Sprint Local Telephone
Companies

BELLSOUTH CORPORATION
BELLSOUTH TELECOMMUNICATIONS, INC.

By: M. Robert Sutherland by JCK

M. Robert Sutherland
Richard M. Sbaratta
Rebecca M. Lough
Suite 1700
1155 Peachtree Street, N.E.
Atlanta, GA 30309-3610
(404) 249-3390

US WEST, INC.

By: Robert B. McKenna by JLK

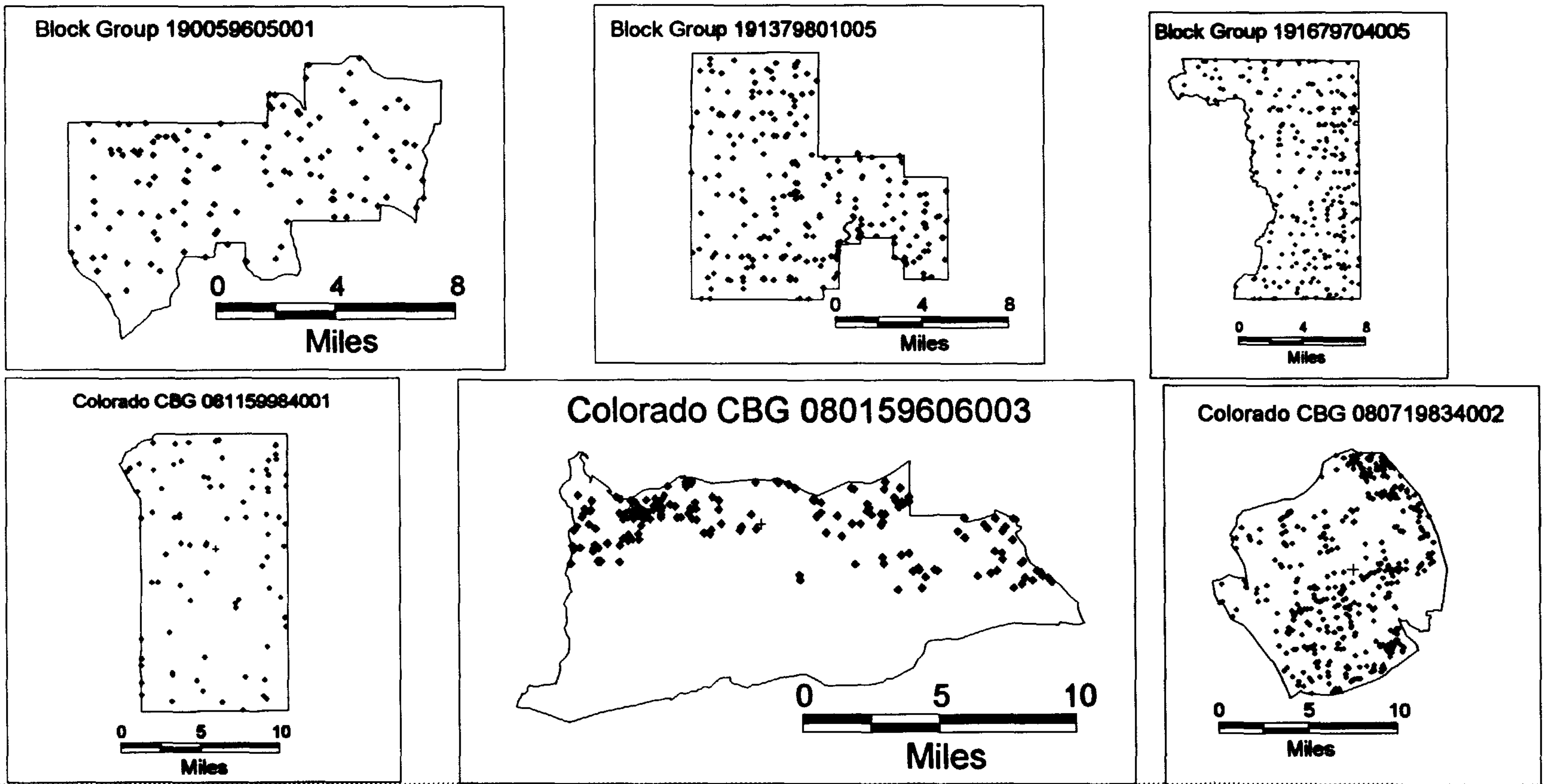
Robert B. McKenna
John L. Traylor
Suite 700
1020 19th Street, N.W.
Washington, D.C. 20036
(303) 672-2798
Of Counsel, Dan L. Poole

ATTACHMENT A

IILUSTRATIVE FIGURES

Figure 1
BCPM Enhanced Customer Location

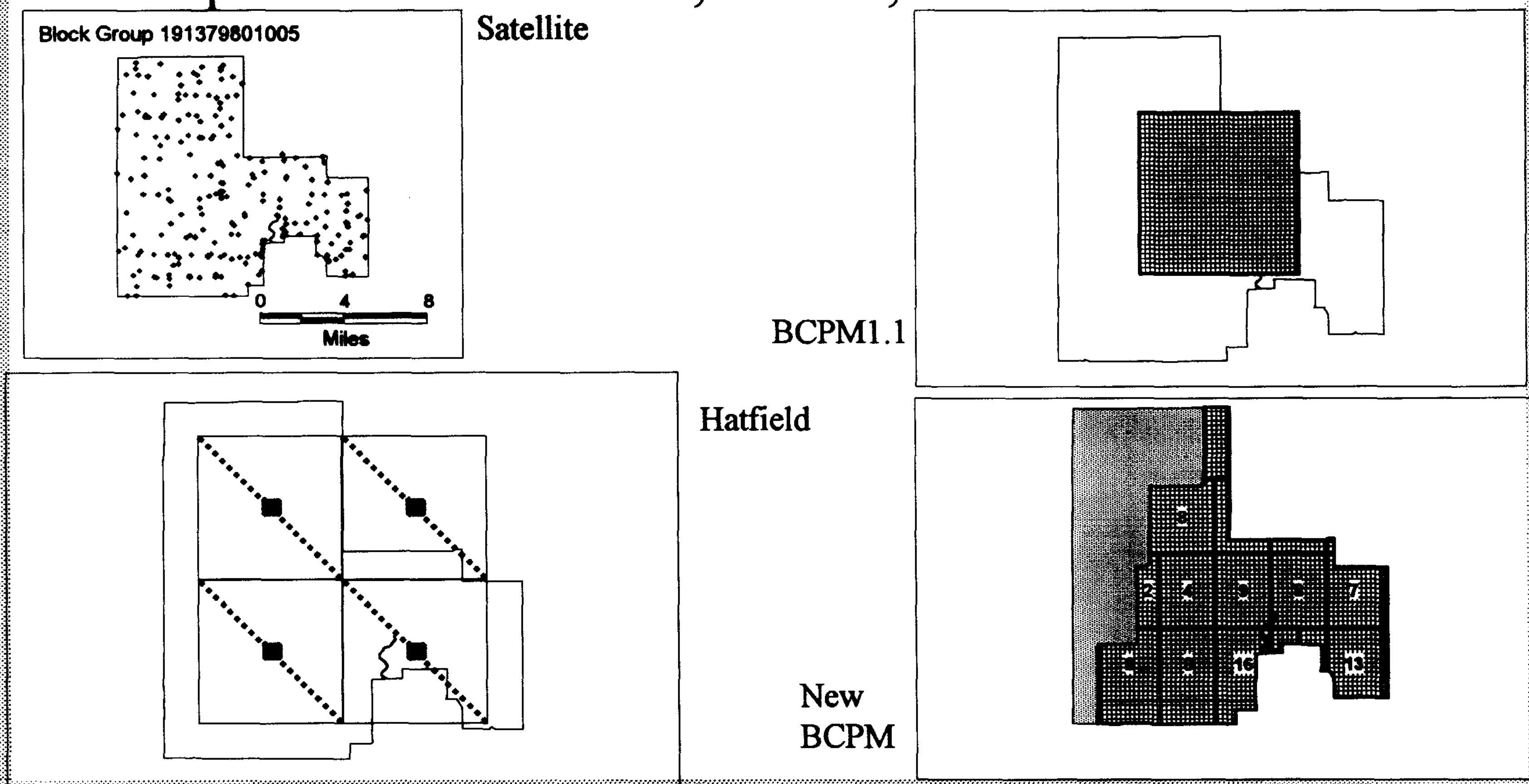
- Digitized Satellite Map Data for Random CBGs with Density < 5/sqmi



Any Representation of data is for illustrative purposes only.

Figure 2
BCPM Enhanced Customer Location

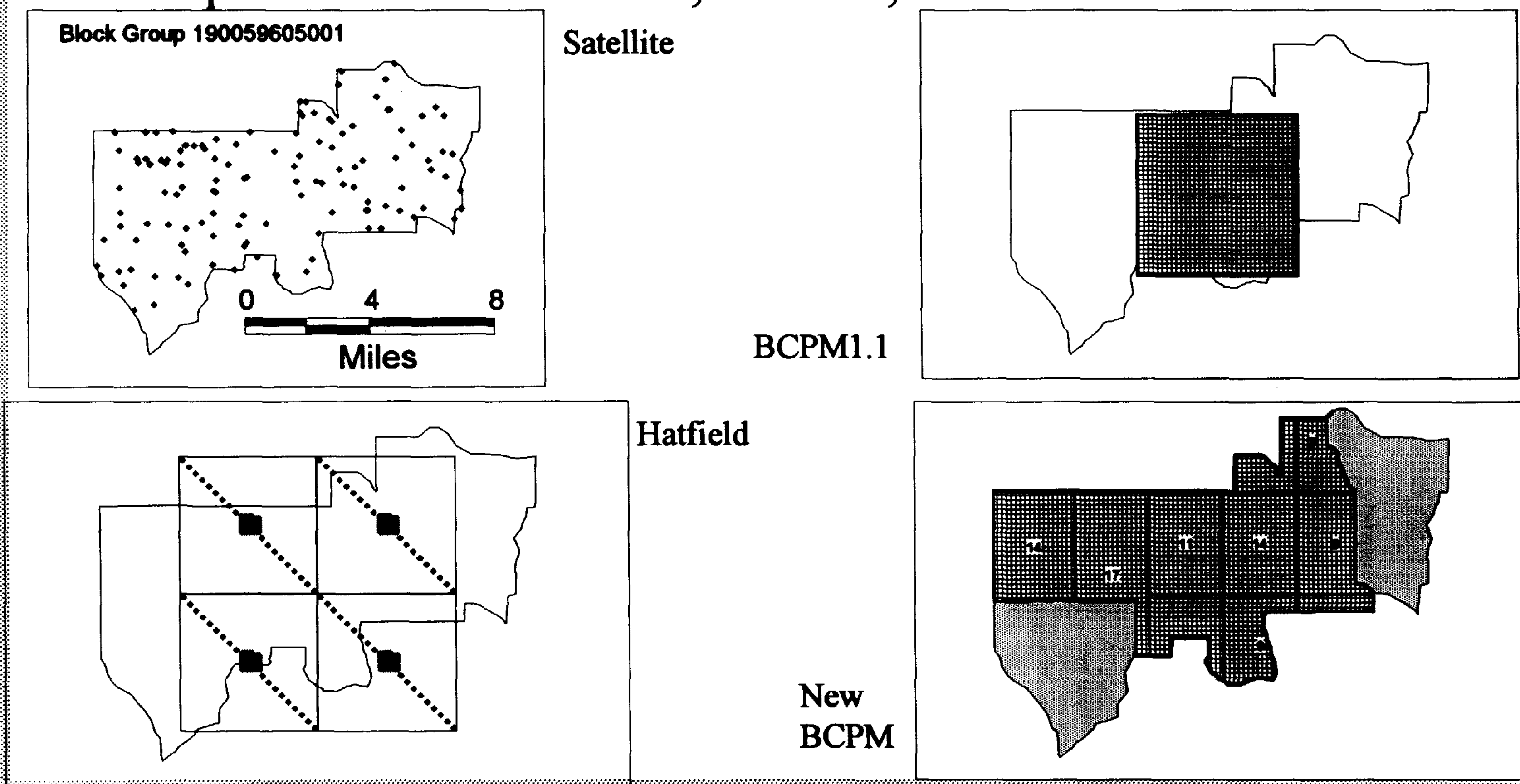
- Comparison of BCPM1.1, HM4.0, And Enhanced BCPM



Any Representation of data is for illustrative purposes only.

Figure 3
BCPM Enhanced Customer Location

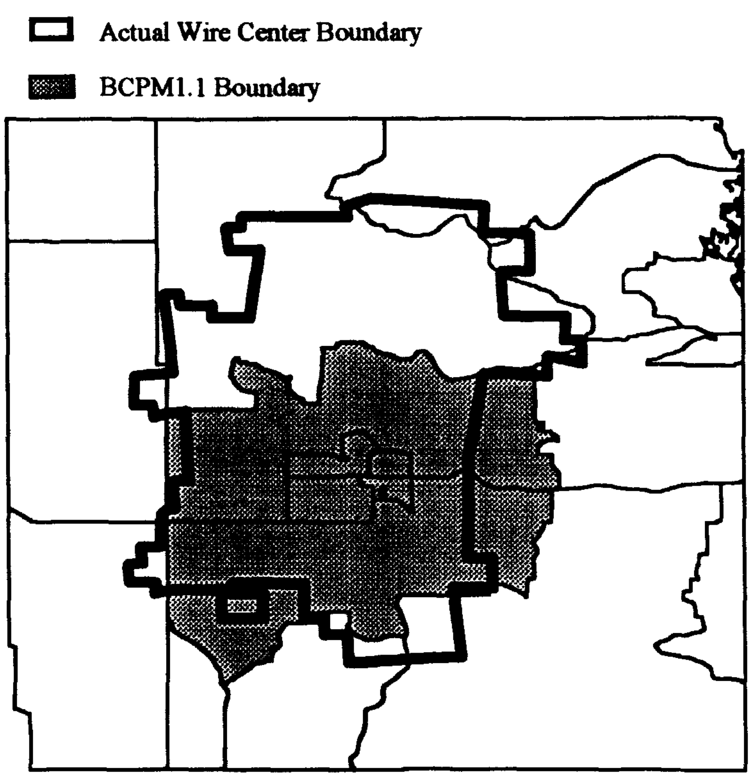
- Comparison of BCPM1.1, HM4.0, And Enhanced BCPM

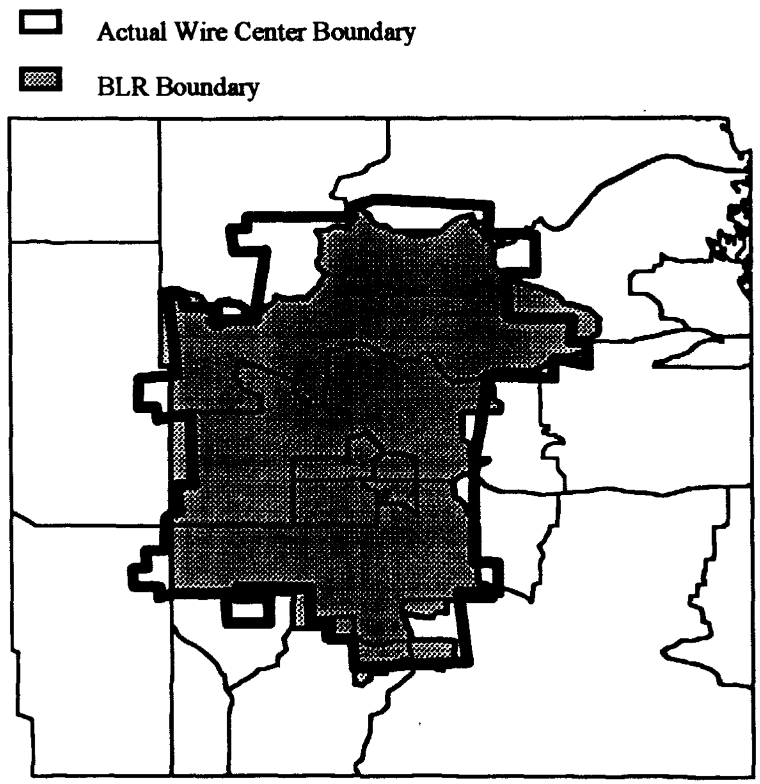


Any Representation of data is for illustrative purposes only.

Figure 4
BCPM Enhanced Customer Location

- Comparison of BCPM1.1, BLR and Actual Wire Center Boundaries

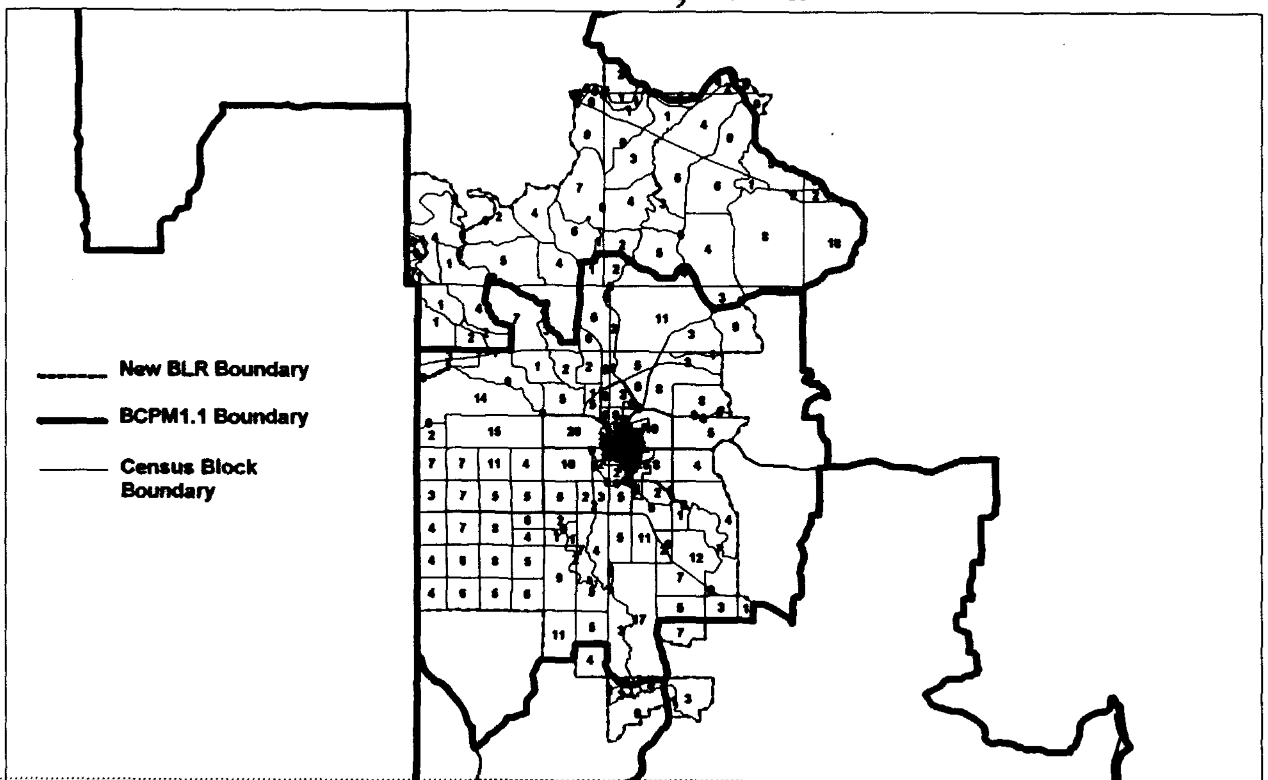




Any Representation of data is for illustrative purposes only.

Figure 5
BCPM Enhanced Customer Location

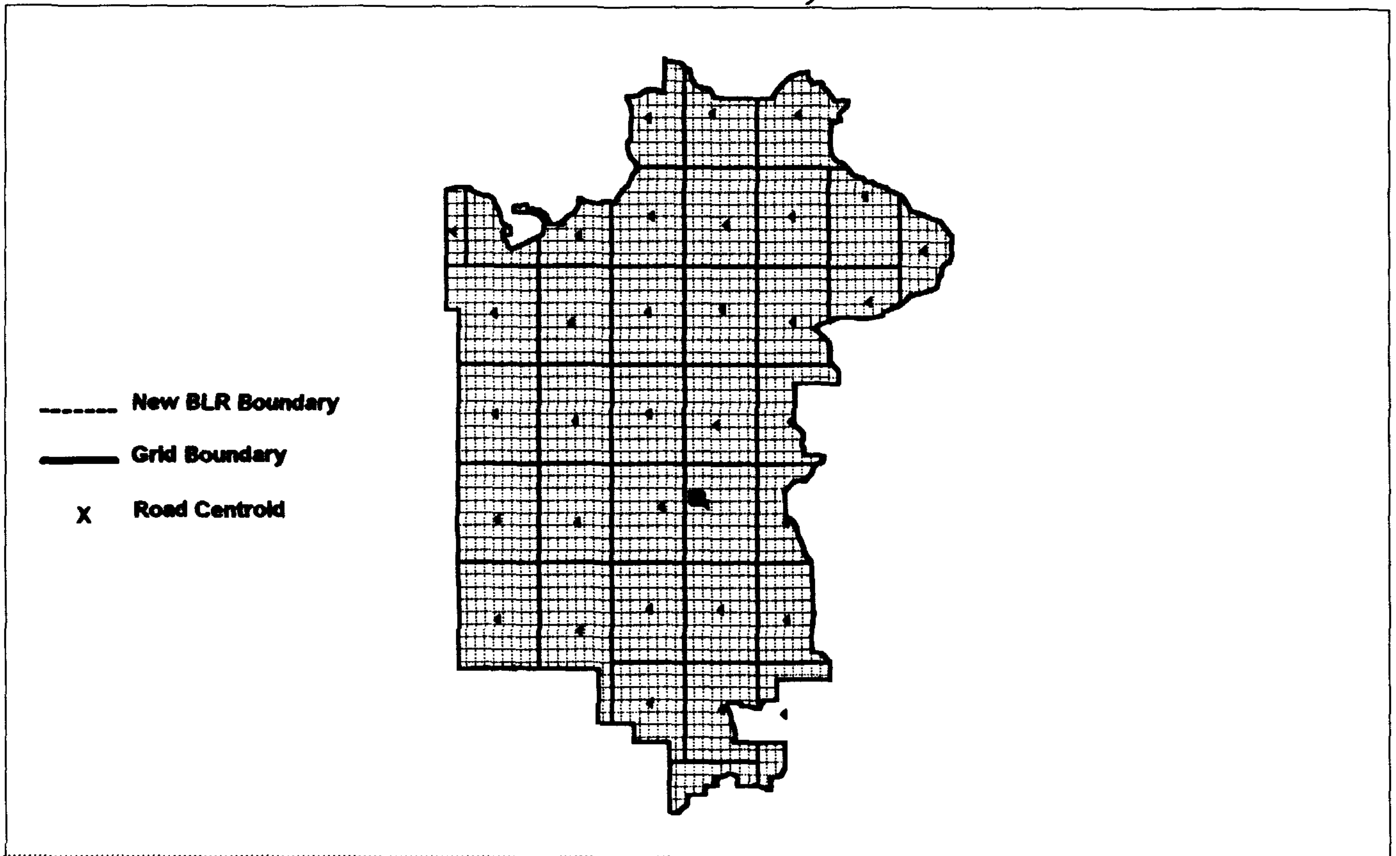
- Finer Level of Input Data
  - Census Block Data for Waukon, Iowa



Any Representation of data is for illustrative purposes only.

Figure 6
BCPM Enhanced Customer Location

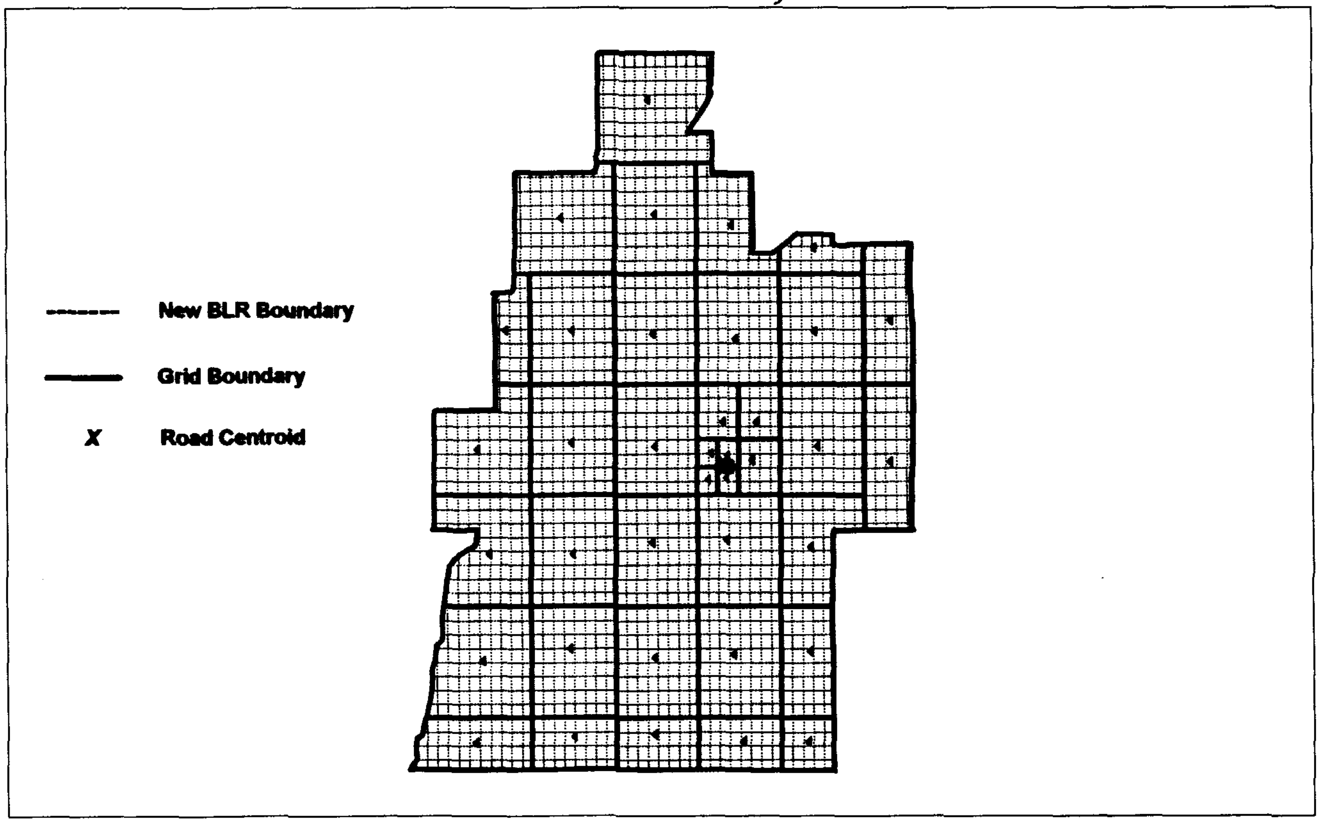
- Variable Size Grids for Waukon, Iowa



Any Representation of data is for illustrative purposes only.

Figure 7
BCPM Enhanced Customer Location

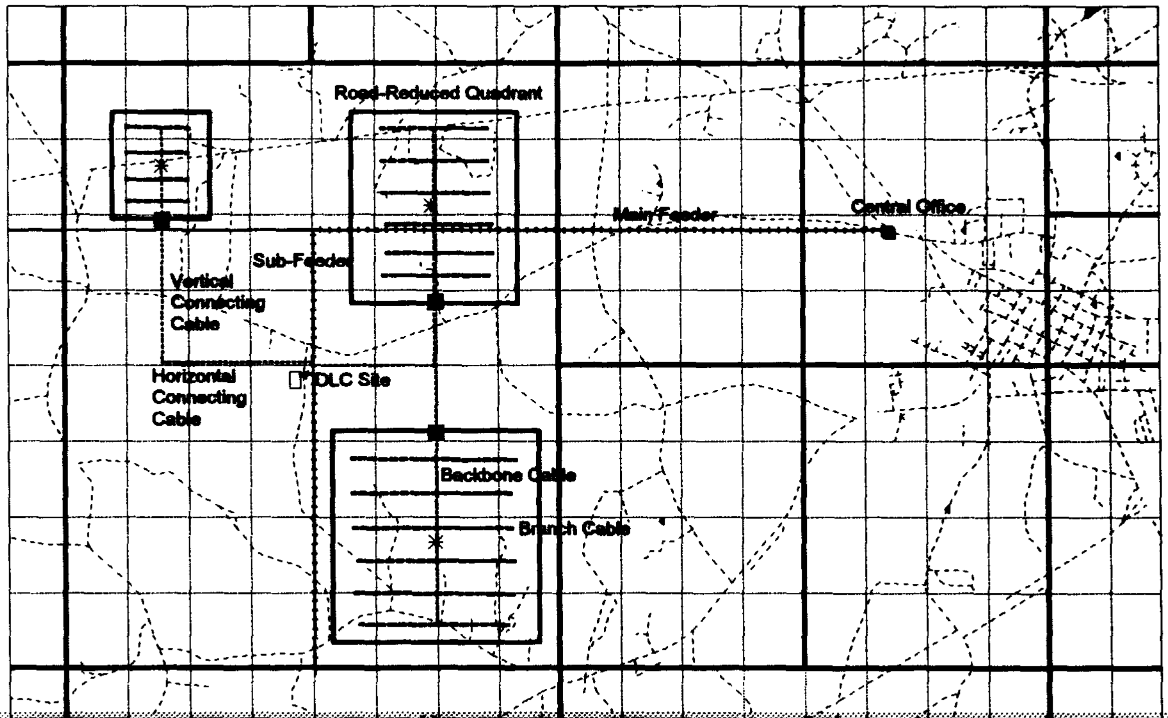
- Variable Size Grids for Red Oak, Iowa



Any Representation of data is for illustrative purposes only.

Figure 8
BCPM Enhanced Customer Location

- Improved Engineering Based Upon Specific Grid Information
  - Grid Engineering: Demonstration of Grid, Quadrants, and Distribution and Feeder Plant



Any Representation of data is for illustrative purposes only.

ATTACHMENT B

An Analysis of the Hatfield 4.0 Model Customer Location Algorithm

September 2, 1997

Audrey J. Curtiss, Ph.D.

Kevin T. Duffy-Deno, Ph.D.

INDETEC International

445 Marine View Ave.
Suite 310
Del Mar, CA 92014
Tel (619) 755-3858
Fax (619) 755-3744
www.INDETEC.com

Table of Contents

I. Introduction

Validation of forward-looking cost proxy models is truly a challenge. The models purport to estimate the cost of building and operating an entirely new wire-based telecommunications network, a network comprised of the most recently available technology. Comparison of this modeled network with existing telephone networks is not straight forward, as various vintages of technology and plant are embedded in companies' cost accounts.

There is one area where a validation check can be made, however. The cost proxy models build a network to serve existing customers in their current locations. Hence, there should be a fairly close correspondence between the physical characteristics of the models' network, characteristics that are spatially dependent and similar to characteristics of exiting networks. In fact, some of the more important drivers of loop cost are the number of lines, the distance of customers from the serving wire center, and the density and dispersion of customers within a company's wire center territory.

The current paper analyzes the Hatfield Model 4.0 on the basis of two reality-based metrics: (1) route or structure miles; and, (2) the level of dispersion in the lowest density areas as implied by the Model's algorithm. The paper also examines the Hatfield Town Factor as it relates to the Hatfield clustering algorithm and demonstrates the error introduced by this algorithm.

This analysis is based on company operating data for the State of Colorado. Specifically, company-level data are used for each of the non-RBOC telephone companies, obtained from the Rural Utilities Service (RUS) and the National Exchange Carriers Association (NECA) for 1995. US Bureau of the Census data are also employed.

The Hatfield 4.0 Model (update 7/31/97) is run for each company in Colorado using the default values for the user-defined inputs. The Hatfield Model now allows the user to specify a Town Factor at the CBG level but does not supply default values at this level of disaggregation. Rather, only the single Town Factor value of 85 % is supplied. To better ascertain the implications of the Hatfield Model's assumptions concerning customer location, a Town Factor value for every CBG in Colorado was separately entered into the

Model. This Town Factor is derived from 1990 US Census data as discussed later in this paper.

II. Two Measures of Model Validation

A. Route Miles of Plant

Route miles of plant refer to the total structure or trench mileage of cable and wire.[10] It represents the extent of a telephone company's network in terms of all the routes required to serve current subscribers. A cost proxy model that purports to accurately estimate the cost of serving existing customers on a geographically disaggregated basis should "build" enough plant to reach all of these customers. Thus, a test of the validity of a cost proxy model is a comparison of the total route miles "built" by the model with actual telephone company data.

Table 1 presents, for the 27 telephone companies operating in Colorado, the total route miles reported by the Rural Utilities Service (RUS) in 1995, and the total route mileage "built" by the Hatfield 4.0 Model. The Hatfield 4.0 Model actually does not provide a calculation of total route miles. The calculation of route miles is made at the CBG level and includes the feeder and subfeeder, the entire length of the vertical and horizontal connecting cables, the total road cable distance for each cluster as well as the backbone and total branch distances in each cluster. The CBG level data are then aggregated to the company level to obtain the total model-estimated route miles shown in Table 1.

Actual route mile data for the Sunflower and Strasburg Telephone companies are estimated as these companies were not part of the RUS loan program in 1995. The estimation technique is the application of a regression-estimated relationship between route miles and cable sheath miles for 16 Colorado telephone companies for which such data are available. Data were obtained from the Colorado Office of Consumer Council. Reported route miles for other companies not part of the RUS loan program, the El Paso County Mutual, Haxtun, Roggen, Pine Drive, Stoneham Cooperative, and Willard Telephone Company, were obtained from the Colorado Office of Consumer Council. These data are considered proprietary, are not shown in Table 1, but were used in the calculation of the average difference between the Model's estimated route miles and reported route miles. At this time, route mile data for the Agate and Phillips County Telephone Company as well as for US West are not available nor is there a feasible method for estimating these data.

As shown in Table 1, the Hatfield Model consistently "under builds" plant. The Model "over builds" plant in the case of only 4 companies.[11] Excluding Eagle Telecommunications, the Hatfield Model under builds plant by an average of 40 %. Assuming that the reported route mile data are accurate, this clearly indicates a deficiency

[10] "Structure" mileage may be a more accurate term since the required calculation reflects buried, underground and aerial cable. However, we will use the more intuitive term "route miles."

[11] The very large discrepancy in the case of Eagle Telecommunications suggests that the 1995 reported data for this company may not be accurate.

on the part of the Hatfield Model to build an adequate network to serve the customers in these locations.

Table 1. Colorado Telephone Companies: Reported Route Miles vs. Model Route Miles

SAR ID	CL	Company	RUS 1995	HM 4.1	% Diff
461835	C	SUNFLOWER TELEPHONE CO.,INC. - CO	249	181	-27.31
462181	C	BIJOU TEL COOPERATIVE ASSOC. INC	374	102	-72.73
462182	C	BLANCA TELEPHONE CO.	303	178	-41.25
462184	C	DELTA COUNTY TELE-COMM INC.	1,050	673	-35.9
462185	C	EAGLE TELECOMMUNICATIONS INC.*	1,293	8,224	536.04
462186	C	EASTERN SLOPE RURAL TEL ASSN INC	2,169	1,222	-43.66
462187	C	EL PASO COUNTY MUTUAL TEL CO		330	
462188	C	FARMERS TEL CO, INC. - COLORADO	221	81	-63.35
462190	C	HAXTUN TELEPHONE COMPANY		495	
462192	C	BIG SANDY TELECOM, INC.	443	258	-41.76
462193	C	NUCLA-NATURITA TEL. CO.	436	287	-34.17
462194	C	NUNN TEL. COMPANY	175	154	-12
462196	C	PEETZ COOP. TEL. CO.	129	54	-58.14
462199	C	PLAINS COOPERATIVE TEL. ASSOC. INC.	1,248	684	-45.19
462201	C	RICO TEL. CO.	33	36	9.09
462202	C	ROGGEN TELEPHONE COOPERATIVE CO.		91	
462203	C	THE RYE TELEPHONE CO. INC.	719	327	-54.52
462204	C	COLUMBINE TELEPHONE COMPANY	432	253	-41.44
462207	C	STRASBURG TEL. CO.	233	245	77.54
462208	C	UNIVERSAL TEL. CO. OF COLORADO	696	617	-11.35
462209	C	WIGGINS TEL. ASSOC.	877	395	-54.96
465102	C	US WEST-COLORADO	NA	33,173	NA
462178	A	AGATE TELEPHONE CO.	NA	113	NA
462197	A	PHILLIPS COUNTY TEL. CO.	NA	149	NA
462198	A	PINE DRIVE TEL. CO.		73	
462206	A	STONEHAM COOPERATIVE TEL. CO.		119	
462210	A	WILLARD TEL. CO.		96	
		Totals Excluding Eagle Telecommunication:			
		With US West		40,124	
		Without US West		6,951	-40.06

One reason why a model may under build plant is if it under-accounts for the number of lines. However, as shown in Table 2, the Hatfield model does include an appropriate number of lines.[12] The actual line count is taken from the year-end working line count reported to the National Exchange Carrier Association (NECA) for 1995. Excluding US West, Table 2 reveals a very close match, on average, between lines reported to NECA and those assumed by the Hatfield Model. The reason for the fairly large discrepancy between line counts for Mountain Bell is the inclusion by the Hatfield Model of a substantial number of special access lines in its modeling of the US West Colorado network.

[12] This paper does not address the issue of the appropriate level of spare capacity or the appropriate difference between number of billed lines and lines constructed.

Another possible reason why the Hatfield Model is under building plant is that it is making an erroneous assumption concerning the dispersion of customers. In other words, if the Model assumes that customers are more clustered than they actually are, then the model will build less plant as compared with the actual network. This is an issue that is particularly relevant for the low-density, rural areas of a state. As the discussion in the following sections reveals, the Hatfield Model does appear to be assuming an unrealistically high degree of clustering in the low-density areas.

Table 2. Colorado Telephone Companies: Actual Lines vs. Model Lines

SAR ID	CL	Company	NECA 1995	HM 4.0	% Diff
461835	C	SUNFLOWER TELEPHONE CO.,INC. - CO	328	336	2.44
462181	C	BIJOU TEL COOPERATIVE ASSOC. INC	1,173	1,161	-1.02
462182	C	BLANCA TELEPHONE CO.	780	797	2.18
462184	C	DELTA COUNTY TELE-COMM INC.	8,013	8,163	1.87
462185	C	EAGLE TELECOMMUNICATIONS INC.*	67,564	67,565	0.00
462186	C	EASTERN SLOPE RURAL TEL ASSN INC	4,542	4,469	-1.61
462187	C	EL PASO COUNTY MUTUAL TEL CO	2.935	2,781	-5.25
462188	C	FARMERS TEL CO, INC. - COLORADO	381	340	-10.76
462190	C	HAXTUN TELEPHONE COMPANY	1,569	1,692	7.84
462192	C	BIG SANDY TELECOM, INC.	843	843	0.00
462193	C	NUCLA-NATURITA TEL. CO.	1,366	1,342	-1.76
462194	C	NUNN TEL. COMPANY	358	357	-0.28
462196	C	PEETZ COOP. TEL. CO.	208	205	-1.44
462199	C	PLAINS COOPERATIVE TEL. ASSOC. INC.	1,510	1,141	-8.36
462201	C	RICO TEL. CO.	137	140	2.19
462202	C	ROGGEN TELEPHONE COOPERATIVE CO.	236	241	2.12
462203	C	THE RYE TELEPHONE CO. INC.	1,684	1,752	4.04
462204	C	COLUMBINE TELEPHONE COMPANY	982	1,010	2.85
462207	C	STRASBURG TEL. CO.	985	1,013	2.84
462208	C	UNIVERSAL TEL. CO. OF COLORADO	6,221	6,573	5.66
462209	C	WIGGINS TEL. ASSOC.	1,403	1,389	-1.00
465102	C	US WEST-COLORADO	2,365,257	2,803,446	18.53
462178	A	AGATE TELEPHONE CO.	109	112	2.75
462197	A	PHILLIPS COUNTY TEL. CO.	1,877	1.936	3.14
462198	A	PINE DRIVE TEL. CO.	637	645	1.26
462206	A	STONEHAM COOPERATIVE TEL. CO.	75	76	1.33
462210	A	WILLARD TEL. CO.	61	61	0.00
		Totals:			
		With US West	2,471,234	2,909,859	17.75
		Without US West	105,977	106,413	0.41

B. *Customer Dispersion*

The preceding route mile analysis suggests that the Hatfield 4.0 Model substantially understates cable investment in the low-density CBG's. This shortcoming of the Hatfield Model may be due in part to its clustering algorithm, which ignores the level of actual customer dispersion within CBG's.

INDETEC International has developed a methodology for comparing the dispersion of housing units that actually exists in a CBG with that implied by the Hatfield 4.0 Model clustering algorithm. The Model's understatement results from its clustering algorithm which forces housing units in a low-density CBG to be located in a very concentrated area. The algorithm does not take into account how the housing units are actually distributed (geographically) within a CBG. That is, the Hatfield clustering algorithm does not take into account the actual degree of housing unit *dispersion*.[13]

The density of housing units in a CBG is *not necessarily* a measure of the dispersion of those units throughout the CBG. For example, a CBG of 100 square miles, occupied by 100 housing units, has a density of one unit per square mile. Dispersion is essentially a measure of the average distance of these units *from one another*. With respect to this imaginary CBG, consider the following two possibilities: (1) nearly all the housing units are clustered together in one compact settlement; and, (2) the housing units are randomly scattered throughout the CBG.

In the first case, with a high degree of clustering, the average distance between housing units is quite small (the dispersion is low). As a result, the length of cable required to serve those units is relatively short on a per subscriber basis. In the second case, with very little clustering, the average distance between housing units is quite large (the dispersion is high). As a result, the length of cable required to serve that same number of units, in a CBG of the same density, is relatively longer than in the clustered case.

The Hatfield 4.0 Model fails to take into account the extent of geographical dispersion of housing units typical in low-density CBG's. The magnitude of this shortcoming is the greatest in the least dense CBG category (less than 5 households per square mile). What follows is an example of this shortcoming based on 1995 CBG data for the US West service territory in the state of Colorado.

Methodology

INDETEC International first developed a methodology for measuring the "actual" dispersion that exists in a given CBG. The ideal measure of dispersion is the average

[13] At paragraphs 222 and 244 in the FCC Universal Service Order discusses "dispersion of loops within a CBG" and "the dispersion of population within a CBG" as issues which must be treated or refined before a proxy model should be used.

distance between housing units in a CBG. However, since information on the exact location of individual units in a CBG is not generally available, INDETEC developed a methodology based on information that is publicly provided by the US Census Bureau at the census block (CB) level.[14] Specifically, the dispersion measure utilized here reflects the average distance from the housing-unit-weighted center of a CBG to the geographic center of each CB in the CBG. This methodology yields a CB-based lower bound for the actual measure of dispersion in a CBG. Any measure of dispersion based on the *actual* locations of all households would result in a higher measure of dispersion.[15]

INDETEC then developed a comparable measure of dispersion based on the Hatfield 4.0 clustering algorithm. Consider any CBG with density less than 200 households per square mile (or where over 50% of the CBG area is unoccupied). The Hatfield 4.0 algorithm overlays, upon the original CBG, a *square* with area equal to that of the occupied Census Blocks (CB's) of the CBG, and models the CBG *as if* it were this square. The algorithm divides the square into four equal square quadrants and establishes a point at the center of each of those quadrants. The Hatfield 4.0 Model then assumes that some user-defined share of the housing units of each quadrant are *clustered* densely about this point (the default value is 85%). It specifically assumes the units are clustered in contiguous 3-acre lots (the 3-acre value is user adjustable). The remaining housing units in the quadrant are uniformly spaced along a line(s) emanating from the center of the clusters ("road cables").

INDETEC developed a measure of quadrant dispersion implied by the Hatfield 4.0 Model algorithm and related it to the dispersion for the CBG as a whole. This Hatfield CBG dispersion measure is then compared to the measure of actual dispersion based on CB data.

Colorado Dispersion Findings: In General

Table 3 shows the dispersion measure for the three least dense zones. Intuitively, one would expect that a measure of dispersion and density be inversely related, holding all else constant.[16] Indeed, this is what Table 3 shows. The weighted average dispersion (weighted by CBG housing units) increases in the lesser dense areas. What is remarkable is that the measure of actual dispersion in the least dense zone is, on average, is 3.5 times that in the next density zone.

Table 3. Colorado Housing Unit Dispersion.

[14] The 226,000 CBG's in the U.S. are comprised of approximately 7 million CB's.

[15] Another measure of dispersion is the average distance between the housing-unit-weighted centers of the CB's. This measure is more costly to produce. However, since INDETEC has determined that it is highly correlated with the CBG center-to-any dispersion measure (0.995), the later measure was chosen for the analysis.

[16] In fact, dispersion is a function of both area and density. A regression of our CBG dispersion measure on occupied area and household density for all CBG's in Colorado with 200 or fewer households per square mile indicates that dispersion is more sensitive to area than to density (elasticities evaluated at the simple sample means are 0.43 for area and –0.18 for density).